

POSTER PRESENTATION

Open Access

# Applying missing data methods to routine data: a prospective, population-based register of people with diabetes

Stephanie Read<sup>1\*</sup>, Sarah Wild<sup>1,2</sup>, Steff Lewis<sup>1</sup>

From 2nd Clinical Trials Methodology Conference: Methodology Matters  
Edinburgh, UK. 18-19 November 2013

## Background

Routinely-collected data could be used to make randomised controlled trials (RCTs) more efficient, either for collection of outcome data or to enhance recruitment. The use of routine data in RCTs has been limited by concerns surrounding data quality, particularly missingness. To exploit these information-rich data sources, it is necessary to identify approaches capable of overcoming high rates of missing data.

## Methods

Using data from a population-based diabetes register linked to mortality records, we compared four methods for handling missing data when investigating the association between body mass index and all-cause mortality in patients with Type 2 diabetes in a retrospective cohort study. Complete case analysis (CCA), population mean imputation (PMI), stochastic imputation (SI) and multiple imputation (MI) methods were applied to handle the missing data. Cox proportional hazard model coefficients for the association between BMI and all-cause mortality were compared for each missing data method.

## Results

Body mass index data were unavailable for 117,048 (54.07%) patients and there were 41,555 deaths among the cohort between 2001 and 2008. Data appeared to be missing at random conditional on year of diagnosis and health status. CCA produced a J-shaped relationship between patient BMI and all-cause mortality, though findings from other approaches indicated that CCA

underestimated the survival in this population. Estimates obtained from SI and MI flattened the observed J-shaped curve. However, imputations were based on poor predictions.

## Summary

Different approaches for handling missing data can influence associations and caution is required when using incomplete routine data to improve RCTs.

## Authors' details

<sup>1</sup>University of Edinburgh, Edinburgh, UK. <sup>2</sup>Scottish Diabetes Research Network Epidemiology Group, Scotland, UK.

Published: 29 November 2013

doi:10.1186/1745-6215-14-S1-P113

**Cite this article as:** Read et al.: Applying missing data methods to routine data: a prospective, population-based register of people with diabetes. *Trials* 2013 **14**(Suppl 1):P113.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)



<sup>1</sup>University of Edinburgh, Edinburgh, UK  
Full list of author information is available at the end of the article